methods

In the report by T.Agarwal it has been stated that there are multiple modelling techniques that can be used to predict energy consumption. The models that suite our needs are listed below:

1. Lineair regression and multiple regression
   As known a lineair regression establishes a relationship between dependent variable (Y) and one or more independent variables(X) using a regression line with an equation such as Y= a+b*x+e . Fitting the regressional line is important, which can be done with the least square method. Which calculates the data by minimizing the sum of squares from each data point to the line.

- There must be **linear relationship** between independent and dependent variables
- Multiple regression suffers from **multicollinearity, autocorrelation, heteroskedasticity**.
- Linear Regression is very sensitive to **Outliers**. It can terribly affect the regression line and eventually the forecasted values.
- Multicollinearity can increase the variance of the coefficient estimates and make the estimates very sensitive to minor changes in the model. The result is that the coefficient estimates are unstable
- In case of multiple independent variables, we can go with **forward selection**, **backward elimination** and **step wise approach** for selection of most significant independent variables.

An example of an algoritm using sklearn:

class `sklearn.linear_model.`**`LinearRegression`**(*fit_intercept=True*, *normalize=False*, *copy_X=True*, *n_jobs=1*)

https://towardsdatascience.com/linear-regression-with-example-8daf6205bd49